Journal of Information Hiding and Multimedia Signal Processing ©2025 ISSN 2073-4212
Ubiquitous International Volume **16**, Number **2**, June **2025**

LoF2ML: Local Feature Selection Based on Few-Shot Multi-Label Deep Learning for Dental Caries Recognition

Son Nguyen Thanh

Department of Economic Information System
Academy of Finance
58 Le Van Hien, Hanoi, Vietnam
sonnt@hvtc.edu.vn

Thanh Nguyen Van*

Department of Economic Information System
Academy of Finance
58 Le Van Hien, Hanoi, Vietnam
nguyenvanthanh@hvtc.edu.vn

*Corresponding author: Thanh Nguyen Van

Received February 5, 2025, revised April 5, 2025, accepted April 8, 2025.

Abstract. *In the problem of dental caries image recognition, our goal is to develop an image recognition approach with limited training data to identify dental caries based on a multi-label and multi-grade recognition framework. A dental caries image typically contains multiple labels and different severity levels, but the available data is quite limited and often contains a significant amount of noise. To address this issue, we propose a method that selectively extracts local features in the multi-label setting from a limited number of dental caries images. This approach helps identify the essential and minimal features necessary to construct a low-sample deep learning model for dental caries image recognition. Experiments conducted on COCO, PASCAL VOC, and T-Caries (our self-collected dental caries dataset) demonstrate that our model significantly outperforms state-of-the-art methods.*
Keywords: Dental caries recognition, few-shot learning, image processing, multi-label learning

---

1. **Introduction.** Identification of dental caries using artificial intelligence (AI) technology [1, 2, 3, 4, 5] is one of the important tasks in today's Industry 4.0 era. Developing machine learning models to assist in recognizing dental caries images can help clinicians reduce diagnosis and treatment time while enhancing the efficiency of medical image interpretation. However, two primary challenges arise when constructing dental caries image recognition systems: (1) Data Scarcity: Dental caries image data are limited and have not been properly collected; (2) Data Labeling: Dental caries images require expert labeling and grading of caries severity by dental professionals, which is both costly and time-consuming. Under these constraints, relying solely on conventional or basic machine learning models for dental caries image recognition is not feasible.

Recently, several deep multi-label learning models [6, 7] have been developed and applied to various image recognition tasks. These models aim to describe an image and identify multiple labels corresponding to different objects or features present in the image. Each label represents a specific task or object. The development of deep multi-label learning techniques aids in automated labeling, data processing, and the enhancement of image feature quality. In this paper, we formulate a multi-label image classification problem for dental caries recognition. Notably, a typical dental image contains various features beyond simply indicating the presence or absence of caries—for example, the severity level of caries or different types of carious lesions.

Another challenge is posed by the dental caries recognition task under conditions of limited training data—referred to as few-shot learning (FSL)—which has attracted significant attention, especially in single-label image recognition problems [8]. These methods primarily rely on exploiting similar features to predict object classes or, at a higher level, on rapidly adapting to new data. However, the accuracy of most FSL approaches remains quite low, particularly for dental caries recognition, where current FSL methods are limited and insufficiently accurate for practical deployment in the medical field. Therefore, we adopt the FSL concept to develop a multi-label FSL approach for dental image recognition that can estimate the regions of the image influencing both the presence of dental caries and its severity level.

In this paper, we propose a model based on a novel similarity measurement method for multi-label dental caries image recognition with limited training data, employing local feature selection (LoF2ML). The selection of local image features allows us to pinpoint specific pixel locations, thereby determining the label for each pixel region. Essentially, we construct groups of embedding vectors to facilitate direct inference, which helps to reduce noise and effectively retrieve image features. We associate these embedding vectors with image features through a dynamic convolution mechanism to incorporate untrained labels without requiring model adjustments. Additionally, a support set is integrated into the model to create a set of categories that the model should focus on, with these categories being directly relevant to the dental caries problem. Given the limited availability of dental caries image data, we also evaluate the proposed model on two public datasets—COCO and PASCAL VOC—and on our own T-Caries dataset (our collected dental caries image dataset).

The main contributions of this work are summarized as follows:

- Develop a few-shot deep learning model based on a multi-label mechanism for dental caries image recognition, named LoF2ML.
- Propose a mechanism that combines embedding vectors with locally selected features to enhance multi-label recognition.
- Construct a dataset related to dental caries images, named T-Caries.
- Evaluate the proposed model on three datasets: COCO, PASCAL VOC, and T-Caries (the dental caries dataset we collected).

2. Related Works.

2.1. Few-shot learning in Dental caries recognition.

Unlike traditional artificial neural network models, low-sample deep learning networks [9, 10, 11] generalize unseen samples by directly learning from a small amount of data. They enhance knowledge in a rich manner by leveraging similarity mechanisms and reusing past knowledge to gradually adapt to untrained data samples. Adaptation capabilities can be implemented through various methods, such as domain adaptation [12], which aims to transfer data acquired from prior experiences to learn new related tasks, or transfer learning [14], which learns based on fully annotated data from related domains to bridge gaps with incomplete data.

Few-shot deep learning proposes a range of improvements to adapt to metadata (meta-learning) [13], enabling training with limited or incomplete data while maintaining the effectiveness of algorithms.

In the task of deep caries image recognition based on deep learning, significant advancements [15] have been made to address practical challenges. The application of few-shot learning models to accurately recognize different caries images with only a small amount of labeled data [13, 9] is bringing new momentum to the field of medical image recognition. Kim [13] proposed a few-shot deep learning model for caries detection on X-ray images. Cherti and colleagues [16] introduced a few-shot deep learning model combined with transfer learning to X-ray image recognition. However, these techniques still rely on feature extraction from entire images using large-scale background data, without focusing on selecting local image features to minimize computational complexity and improve processing speed.

2.2. **Multi-label Few-shot learning in Dental caries recognition.** Labeled image recognition (Multi-label recognition) [17, 18] has attracted significant attention from researchers in recent years. Various solutions have been proposed to build learning capabilities using binary classifiers for recognizing a list of categories [19]; however, a major drawback is the lack of consideration for label correlations. Some attention-based networks [20, 21] have been introduced to integrate multi-label recognition while capturing label dependencies. Sehar et al. [22] utilized deep learning to detect caries in wisdom teeth. Lee et al. [23] investigated deep learning for identifying and classifying the severity of dental caries in dental X-ray images. These models aim to develop caries recognition based on images, striving to identify as many caries labels or severity levels as possible. However, most approaches tend to focus on single-label recognition based on probability estimations, without fully leveraging multi-label image information.

There have been few studies exploring both multi-label recognition and training few-shot deep learning models for dental caries image recognition. Therefore, developing a few-shot deep learning model based on multi-label recognition [24, 25] presents both a challenge and an opportunity for implementation in the medical field. Moukheiber et al. [25] proposed a few-shot deep learning model based on multi-label image recognition for classifying chest X-ray images. Wang et al. [26] applied few-shot deep learning with multi-label classification for lung cancer detection, utilizing vision transformers for feature extraction. The combination of few-shot deep learning and multi-label image recognition is a relatively new trend, and there are still many limitations, particularly in data processing and accuracy.

3. Proposed Method.

3.1. **Problem Definition.** We formulate dental caries recognition as a local feature selection based on multi-label few-shot image classification (LoF2ML) task. Let $\mathcal{C}_{\text{base}}$ and $\mathcal{C}_{\text{novel}}$ denote disjoint sets of base classes (e.g., known caries types: occlusal, root caries; severity levels: mild, moderate) and novel classes (e.g., emerging patterns like enamel lesions), respectively. Given a base dataset $\mathcal{E}_{\text{base}}$ (annotated with $\mathcal{C}_{\text{base}}$) and a novel dataset $\mathcal{E}_{\text{novel}}$ (annotated with $\mathcal{C}_{\text{novel}}$), the goal is to train a model on $\mathcal{E}_{\text{base}}$ to generalize to $\mathcal{C}_{\text{novel}}$ using only K-shot examples. Each image may exhibit *multiple labels* (e.g., "occlusal caries + severe"), necessitating local feature analysis to resolve fine-grained ambiguities.

To capture discriminative regions (e.g., lesion boundaries, demineralized areas), we extract local features $\mathbf{F} = \{f_1, f_2, \ldots, f_N\}$ from N spatial regions $\{R_1, R_2, \ldots, R_N\}$ of an

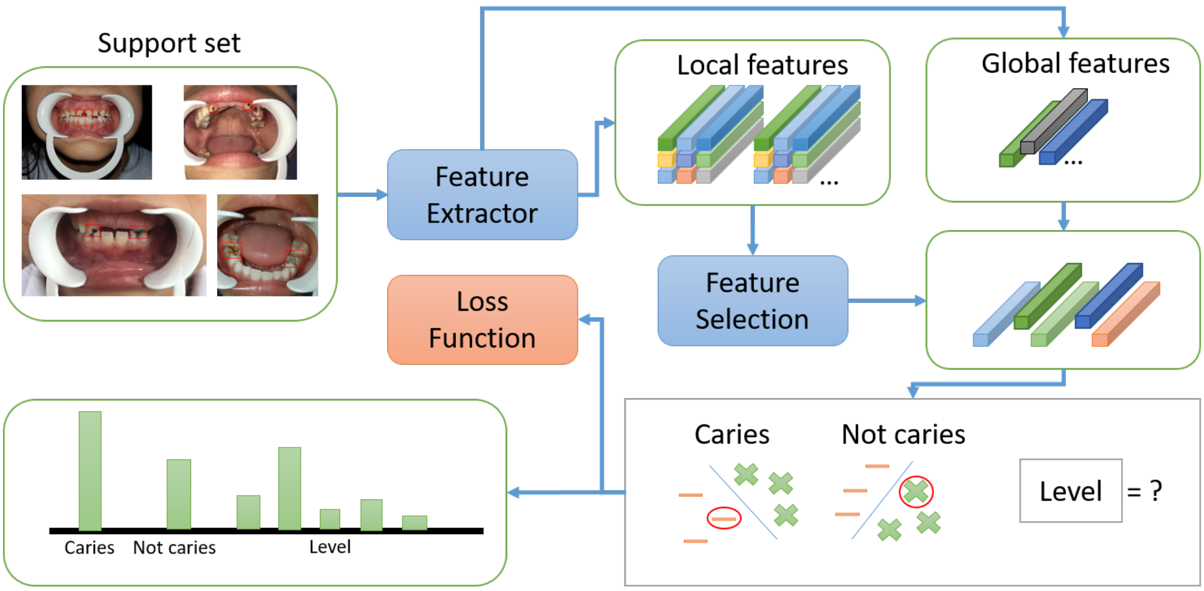

FIGURE 1. The framework of the proposed LoF2ML

image $\mathbf{I}$. An attention mechanism assigns weights α_i to each region:

$$\alpha_i = \frac{\exp\left(\phi(f_i, \mathbf{W}_a)\right)}{\sum_{j=1}^{N} \exp\left(\phi(f_j, \mathbf{W}_a)\right)}, \tag{1}$$

where $\phi(\cdot)$ is a learnable function parameterized by $\mathbf{W}_a$. The aggregated feature $\tilde{\mathbf{F}} = \sum_{i=1}^{N} \alpha_i f_i$ prioritizes caries-specific patterns, enabling interpretable focus on clinically critical regions. Training follows an episodic paradigm: for each episode, a support set $\mathcal{S}$ (containing K images per class from $\mathcal{C}_{\text{base}}$) and a query set $\mathcal{Q}$ (sampled similarly) guide the model to learn transferable representations for multi-label prediction. Key challenges include (1) *multi-label co-occurrence*, where overlapping labels confuse feature attribution; (2) *few-shot instability*, as limited novel examples exacerbate class imbalance; and (3) *local feature ambiguity*, where subtle lesions correlate with multiple labels. This framework supports clinics in diagnosing rare caries types with minimal data, leveraging interpretable attention maps to highlight diagnostically critical regions.

3.2. **Model Architecture.** The LoF2ML (Local Feature-based Few-shot Multi-label Learning) model is designed to address the problem of dental caries detection by leveraging local feature extraction and the few-shot learning mechanism, show in Figure 1. The model architecture consists of key components: a feature extractor, separation of features into local and global categories, feature selection, and multi-label classification.

The model's input is an image of the patient's oral cavity, in which carious regions are marked. The input image is processed through a Feature Extractor, which transforms visual information into machine-processable feature representations. This extractor employs a deep learning network to generate two types of features: Local Features and Global Features. Local Features focus on critical regions of the image, such as carious lesions, color variations, and structural changes in the teeth, while Global Features capture general characteristics of the oral cavity and dental structure.

After feature extraction, these features are passed through the Feature Selection Module, which eliminates irrelevant information and retains only the most influential factors for diagnosis. This feature selection step enables the model to focus more effectively on regions associated with dental caries and the severity of the damage.

Following the selection process, the model utilizes a Few-shot Learning Mechanism to aggregate information from a limited number of labeled images. This approach enhances system performance even when the training dataset is constrained. Additionally, the model incorporates a Multi-label Classification mechanism, enabling it to classify images into different categories, such as "caries present" or "caries absent", while also determining the severity of the lesion.

Another crucial component of the model is the Loss Function, which optimizes the learning process by adjusting weights to ensure accurate label classification. Once training is complete, the model performs predictions by leveraging the selected features to make a final decision on the patient's dental caries status. With this architecture, LoF2ML achieves accurate dental caries even with limited training data while effectively utilizing key features to enhance detection performance.

3.3. **Loss function.** Initially, an input image of a patient's oral cavity is processed by the feature extractor, which utilizes a deep neural network to generate two types of feature representations: local features ($\mathbf{F}_l$) and global features ($\mathbf{F}_g$). The local features capture fine-grained details such as lesion areas, discoloration, and structural abnormalities, whereas global features provide a holistic representation of the oral cavity and overall dental structure. These extracted features are represented as a set of vectors in a shared embedding space:

$$\mathbf{F} = \{f_1, f_2, \ldots, f_N\}, \quad f_i \in \mathbb{R}^d \tag{2}$$

where N is the number of extracted feature regions, and d denotes the dimensionality of the feature vectors.

Following feature extraction, the model employs a feature selection mechanism to eliminate irrelevant features and retain only the most relevant attributes for caries detection. This process involves selecting the most important local features based on their cosine similarity to a reference embedding vector $\mathbf{w}$, which serves as a prototype for dental caries patterns:

$$\mathbf{u}_i = \arg\max_{\mathbf{F}_l} \cos(\mathbf{F}_l, \mathbf{w}) \tag{3}$$

To refine the selected features, two sets of dynamically generated convolutional kernels are derived from $\mathbf{w}$ using linear transformation layers:

$$\mathbf{k}_1 = \mathbf{W}_1\mathbf{w}, \quad \mathbf{k}_2 = \mathbf{W}_2\mathbf{w} \tag{4}$$

where $\mathbf{W}_1$ and $\mathbf{W}_2$ are trainable weight matrices. These convolutional kernels are subsequently applied in a two-step process for refining the feature selection.

The first step involves nonlinear activation and normalization, where a ReLU activation function and Layer Normalization are used to enhance feature separability:

$$\mathbf{u} = \mathrm{ReLU}(\mathrm{Norm}(\mathbf{k}_1\mathbf{u})) \tag{5}$$

The second step computes the category prototype, which provides an optimal representation for each caries class by aggregating refined local feature embeddings:

$$\mathbf{p} = \sum_{i=1}^{N} \mathrm{ReLU}(\mathrm{Norm}(\mathbf{k}_2\mathbf{u}_i)) \tag{6}$$

This category prototype $\mathbf{p}$ enables the model to identify regions of interest within the image that correspond to different levels of dental caries severity.

Once the feature selection process is complete, the model incorporates few-shot learning (FSL) to generalize across limited labeled samples. Additionally, the multi-label classification module enables the system to classify images into multiple categories, such as "caries" or "non-caries," while simultaneously predicting the severity level of detected caries.

To optimize the learning process, the model employs a composite loss function that combines a cross-entropy classification loss with a similarity-based loss to improve feature alignment:

$$\mathcal{L} = \mathcal{L}_{\text{CE}} + \mathcal{L}_{\text{Sim}} \tag{7}$$

where $\mathcal{L}_{\text{CE}}$ is the cross-entropy loss for classification, and $\mathcal{L}_{\text{Sim}}$ ensures similarity between selected features.

With this architecture, LoF2ML can accurately detect dental caries while effectively handling limited training data. The dynamic feature selection mechanism enables the model to focus on the most critical regions, reducing noise and improving classification performance.

4. **Experiments.**

4.1. **Dataset.**

We conducted experiments on three different datasets: COCO, Pascal VOC, and T-Caries.

First, the COCO dataset [27] was used to evaluate the FSL model in a previous study, where the dataset was split into 64 training labels and 16 test labels. However, since this split did not include a validation set, we restructured the 64 training labels into 12 labels for validation (including objects such as cow, dining table, zebra, sandwich, bear, toaster, person, laptop, bed, teddy bear, baseball bat, snowboard) and 52 labels for training, while keeping the 16 test labels unchanged. We followed the approach in [27] by utilizing images from the training and validation sets of COCO 2014. Images that did not contain any test or validation labels were used for training, and the validation set contained only images without training or test labels.

Second, we proposed a new dataset based on PASCAL VOC [28], consisting of 20 labels. To maximize the number of images, we selected six labels for the new Cnovel set (dog, sofa, cat, potted plant, TV monitor, sheep) and six other labels for the validation set (boat, cow, train, airplane, bus, bird). The remaining eight labels were used for training. We used images from the training, validation, and test sets of VOC 2007, as well as the training and validation sets of VOC 2012 (noting that the labels for the VOC 2012 test set are not publicly available).

Finally, the T-Caries dataset comprises 2,440 images labeled with the presence and severity of dental caries. Each image may contain multiple bounding boxes indicating the location of caries with different severity levels. For images with caries, each lesion is further annotated with one of five severity levels (Level 1 to Level 5), representing increasing degrees of carious progression. An image may contain multiple carious lesions, each marked with a bounding box and an associated severity label. The dataset includes approximately 200 to 300 annotated samples for each severity level, ensuring a relatively balanced distribution across classes. All annotations were performed by qualified dental professionals following standardized diagnostic guidelines. This detailed labeling makes the dataset a valuable resource for training and evaluating deep learning models for automatic caries detection and severity classification. This dataset serves as a valuable resource for researching and developing deep learning models in dentistry, particularly for automatic caries detection and classification.

TABLE 1. Experimental results (F1-Score) on COCO, Pascal VOC and T-Caries dataset

Method	COCO		Pascal VOC		T-Caries	
	1-shot	5-shot	1-shot	5-shot	1-shot	5-shot
Prototypical Network [29]	0.380	0.512	0.593	0.652	0.482	0.536
Relation Network [30]	0.395	0.514	0.615	0.663	0.475	0.523
CMNet [31]	0.312	0.331	0.581	0.599	0.397	0.438
RePRI [32]	0.330	0.436	0.475	0.603	0.381	0.429
NTRENet [33]	0.381	0.417	0.583	0.668	0.466	0.495
CyCTR [34]	0.398	0.470	0.590	0.650	0.478	0.528
LoF2ML (ours)	**0.404**	**0.523**	**0.637**	**0.683**	**0.518**	**0.554**

4.2. **Experiment setup.** The LoF2ML model was trained for 100 epochs using the Adam optimizer with a learning rate of 0.0001. During the testing phase, we evaluated the model using 150 test episodes. The ResNet-100 backbone was employed as the feature extractor, with input dimensions resized to 512×512. Additionally, we leveraged an early stopping mechanism at epoch 68 to mitigate overfitting issues.

Our experiments were conducted to address two key research questions (RQs):

- **RQ1:** How does the LoF2ML model perform compared to state-of-the-art models within the same concept?
- **RQ2:** How well does the LoF2ML model predict in real-world dental caries image recognition tasks?

4.3. **Performance Compare (RQ1).** Table 1 presents the F1-score values of various methods on three datasets: COCO, Pascal VOC, and T-Caries, under both 1-shot and 5-shot settings. The proposed method, LoF2ML, achieves the highest performance across all datasets and experimental conditions, demonstrating its stability, generalization capability, and effectiveness in handling multi-label classification tasks in low-sample learning scenarios.

On the COCO dataset, LoF2ML achieves an F1-score of 0.404 (1-shot) and 0.523 (5-shot), outperforming the second-best method, CyCTR, with an improvement of 1.5% (1-shot) and 5.3% (5-shot). This result demonstrates LoF2ML's ability to effectively handle complex multi-label classification tasks involving diverse object categories. Similarly, on the Pascal VOC dataset, LoF2ML maintains its lead with an F1-score of 0.637 (1-shot) and 0.683 (5-shot), surpassing Relation Network by 3.6% and 3.0%, respectively. This highlights the effectiveness of leveraging local features and relational information between labels.

Notably, on the T-Caries dataset, LoF2ML achieves 0.518 (1-shot) and 0.554 (5-shot), improving by 7.5% (1-shot) and 3.4% (5-shot) over the second-best method, Prototypical Network. This indicates that LoF2ML extracts local features more effectively, making it particularly well-suited for medical lesion detection, where identifying small details is crucial.

A detailed analysis of individual methods reveals that Prototypical Network performs well on Pascal VOC and T-Caries but struggles on COCO due to its heavy reliance on global features, limiting its ability to handle datasets with diverse labels. Relation Network achieves good results on Pascal VOC but experiences a performance drop on T-Caries, indicating limitations in processing medical data with fine-grained details. Meanwhile, CMNet and RePRI show the lowest performance across all three datasets, reflecting the inadequacy of their architectures for multi-label classification in few-shot learning

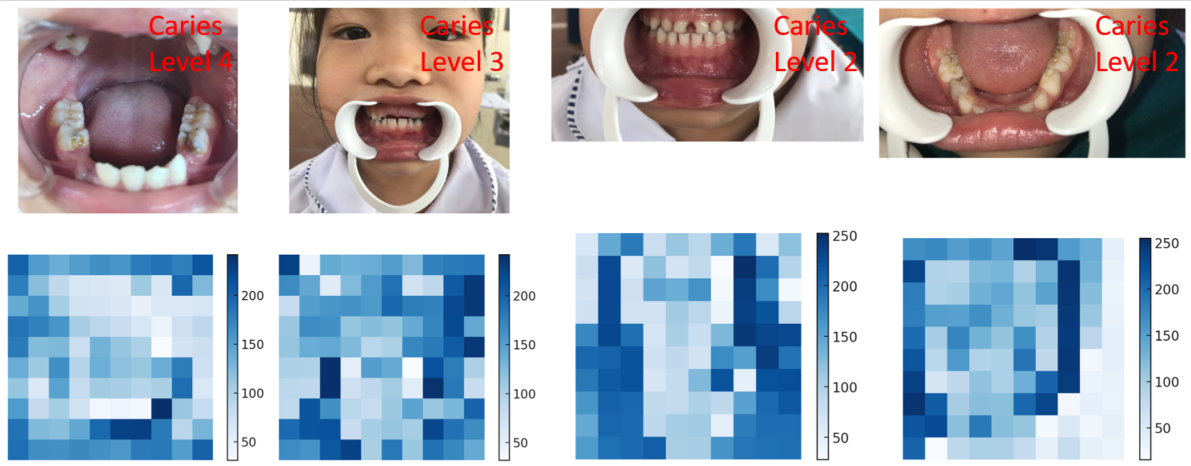

FIGURE 2. Visualization of attention weights between local features is used for deep learning of few-pattern dental caries recognition.

scenarios. CyCTR performs well on COCO but declines significantly on T-Caries, highlighting the challenges of generalist approaches when applied to specialized tasks.

Additionally, the results show that all methods improve when the number of training samples increases from 1-shot to 5-shot, confirming that additional training data significantly enhances model performance. The performance gap between LoF2ML and other methods is most pronounced on COCO and T-Caries, indicating that LoF2ML excels in scenarios involving diverse labels (COCO) and dental caries detection (T-Caries).

4.4. **Qualitative Study (RQ2).** Figrure 2 above illustrates the process of identifying and classifying the severity of dental caries by combining clinical images with feature maps, reflecting the deep learning capability of the LoF2ML model (Local Feature-based Few-shot Multi-label Learning). Five levels of caries are classified from Caries Level 1 to Caries Level 5, corresponding to increasing degrees of damage. Caries Level 1 typically manifests as white spots or mild demineralization areas on the enamel, without obvious damage. Caries Level 2 involves the appearance of small cavities on the enamel, but the dentin is not yet affected. By Caries Level 3, the damage spreads, penetrating the dentin, causing larger areas of demineralization and potentially leading to pain. Caries Level 4 shows severe tooth tissue loss, exposed dentin, and a risk of pulp inflammation, while Caries Level 5 is the most severe stage, where the tooth is completely destroyed, possibly accompanied by abscesses, and requires restorative treatment or extraction.

Beneath the clinical images, feature maps utilize an attention mechanism to highlight the most critical lesion areas in each image. The dark blue regions represent the areas where the model focuses the most, typically the locations of dental caries. At Caries Level 1 and 2, the feature maps show a relatively scattered distribution, reflecting mild damage. In contrast, Caries Level 3, 4, and 5 exhibit higher density concentrations, with lesion areas clearly identified by the model. This demonstrates the LoF2ML model's ability not only to detect dental caries but also to determine the severity of the condition based on image features.

The combination of clinical images and feature maps proves the model's effectiveness in few-shot learning, enabling the system to classify accurately even with limited training data. The model can address multi-label classification problems, where an image may contain multiple different features, such as lesions on multiple teeth or various levels within the same oral cavity. The Feature Selection Module helps the model focus on critical regions, eliminate noise, and optimize diagnostic accuracy.

5. **Conclusions.** We have developed a few-shot deep learning model combined with multi-label classification to predict and identify dental caries images. The primary objective is to selectively extract local features to avoid data redundancy and enhance the efficiency of the image recognition task. Through both quantitative and qualitative evaluations on three datasets—COCO, PASCAL VOC, and T-Caries—the proposed model has demonstrated its effectiveness. However, the accuracy in dental caries image recognition remains limited due to the complexity of the data. Therefore, in the future, we aim to improve the feature extraction process and strive to achieve real-time prediction capabilities.

REFERENCES

[1] Kühnisch, J., et al. "Caries detection on intraoral images using artificial intelligence." Journal of dental research 101.2 (2022): 158-165.

[2] Mertens, Sarah, et al. "Artificial intelligence for caries detection: randomized trial." Journal of dentistry 115 (2021): 103849.

[3] Budi Setiyono, Dwi Ratna Sulistyaningrum, Ario Fajar Pratama, Ridho Nur Rohman Wijaya "A Modification of the Temporal Group Attention Method on Super-Resolution Video for Vehicle Number Plate Detection" Journal of Information Hiding and Multimedia Signal Processing, Vol. 15, No. 3, pp. 156-165, September 2024.

[4] Jasmy Davies, Sivakumari S "A Comparative Analysis of Destructive Methods and Non-Destructive Methods with Machine Learning and Deep Learning Approaches for Rice Leaf Disease Identification" Journal of Information Hiding and Multimedia Signal Processing, Vol. 15, No. 2, pp. 87-97, June 2024.

[5] Tran-Anh, Dat, et al. "Integrative zero-shot learning for fruit recognition." Multimedia Tools and Applications 83.29 (2024): 73191-73213.

[6] Tarekegn, Adane Nega, Mohib Ullah, and Faouzi Alaya Cheikh. "Deep learning for multi-label learning: A comprehensive survey." arXiv preprint arXiv:2401.16549 (2024).

[7] Liu, Weiwei, et al. "The emerging trends of multi-label learning." IEEE transactions on pattern analysis and machine intelligence 44.11 (2021): 7955-7974.

[8] Kim, Chae-Lin, and Byung-Gyu Kim. "Few-shot learning for facial expression recognition: a comprehensive survey." Journal of Real-Time Image Processing 20.3 (2023): 52.

[9] Park, Eun Young, et al. "Caries detection with tooth surface segmentation on intraoral photographic images using deep learning." BMC Oral Health 22.1 (2022): 573.

[10] Cao, Xiaofeng, et al. "A survey of learning on small data: generalization, optimization, and challenge." arXiv preprint arXiv:2207.14443 (2022).

[11] Gong, Yunye, et al. "Recovering manifold representations via unsupervised meta-learning." Frontiers in Computer Science 6 (2025): 1255517.

[12] Wang, Fan, et al. "Exploring domain-invariant parameters for source free domain adaptation." Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2022.

[13] Kim, Min Joo, et al. "Unsupervised few shot learning architecture for diagnosis of periodontal disease in dental panoramic radiographs." Scientific Reports 14.1 (2024): 23237.

[14] Alsakar, Yasmin M., et al. "Multi-label dental disorder diagnosis based on MobileNetV2 and swin transformer using bagging ensemble classifier." Scientific Reports 14.1 (2024): 25193.

[15] Lee, Jae-Hong, et al. "Detection and diagnosis of dental caries using a deep learning-based convolutional neural network algorithm." Journal of dentistry 77 (2018): 106-111.

[16] Cherti, Mehdi, and Jenia Jitsev. "Effect of pre-training scale on intra-and inter-domain, full and few-shot transfer learning for natural and X-Ray chest images." 2022 International Joint Conference on Neural Networks (IJCNN). IEEE, 2022.

[17] Guo, Zixian, et al. "Texts as images in prompt tuning for multi-label image recognition." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2023.

[18] Cole, Elijah, et al. "Multi-label learning from single positive labels." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2021.

[19] Akinola, Olatunji O., et al. "Multiclass feature selection with metaheuristic optimization algorithms: a review." Neural Computing and Applications 34.22 (2022): 19751-19790.

[20] Yan, Jiaxing, et al. "Channel-wise attention-based network for self-supervised monocular depth estimation." 2021 International Conference on 3D vision (3DV). IEEE, 2021.

[21] Zhou, Wei, et al. "Double attention based on graph attention network for image multi-label classification." ACM Transactions on Multimedia Computing, Communications and Applications 19.1 (2023): 1-23.
[22] Sehar, Uroosa, Jing Xiong, and Zeyang Xia. "Automatic Tooth labeling After Segmentation using Prototype-based Meta-learning." Machine Intelligence Research (2025).
[23] Lee, Shinae, et al. "Deep learning for early dental caries detection in bitewing radiographs." Scientific reports 11.1 (2021): 16807.
[24] Hou, Yutai, et al. "Few-shot learning for multi-label intent detection." Proceedings of the AAAI Conference on Artificial Intelligence. Vol. 35. No. 14. 2021.
[25] Moukheiber, Dana, et al. "Few-shot learning geometric ensemble for multi-label classification of chest x-rays." MICCAI Workshop on Data Augmentation, Labelling, and Imperfections. Cham: Springer Nature Switzerland, 2022.
[26] Wang, Chong, et al. "Cross-and Intra-image Prototypical Learning for Multi-label Disease Diagnosis and Interpretation." arXiv preprint arXiv:2411.04607 (2024).
[27] Jain, Swasti, Sonali Dash, and Rajesh Deorari. "Object detection using coco dataset." 2022 International Conference on Cyber Resilience (ICCR). IEEE, 2022.
[28] Tong, Kang, and Yiquan Wu. "Rethinking PASCAL-VOC and MS-COCO dataset for small object detection." Journal of Visual Communication and Image Representation 93 (2023): 103830.
[29] Laenen, Steinar, and Luca Bertinetto. "On episodes, prototypical networks, and few-shot learning." Advances in Neural Information Processing Systems 34 (2021): 24581-24592.
[30] Li, Zijun, et al. "SaberNet: Self-attention based effective relation network for few-shot learning." Pattern Recognition 133 (2023): 109024.
[31] Lai, Lvlong, et al. "CMNet: Component-Aware Matching Network for Few-Shot Point Cloud Classification." IEEE Transactions on Multimedia (2024).
[32] Boudiaf, Malik, et al. "Few-shot segmentation without meta-learning: A good transductive inference is all you need?." Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2021.
[33] Liu, Yuanwei, et al. "Learning non-target knowledge for few-shot semantic segmentation." Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2022.
[34] Zhang, Gengwei, et al. "Few-shot segmentation via cycle-consistent transformer." Advances in Neural Information Processing Systems 34 (2021): 21984-21996.